

ED 015 671

EM 005 629

SUMMARY OF RESEARCH IN EXPERIMENTAL EDUCATION.

BY- BERMAN, MARK LAURENCE
ARIZONA STATE UNIV., TEMPE

PUB DATE AUG 67

EDRS PRICE MF-\$0.25 HC-\$0.80 18P.

DESCRIPTORS- *PROGRAMED INSTRUCTION, *PERFORMANCE FACTORS,
*CULTURALLY DISADVANTAGED, *LEARNING DIFFICULTIES, *DROPOUTS,
MEXICAN AMERICANS, AMERICAN INDIANS, ADULTS, ELEMENTARY
SCHOOL STUDENTS

THE AUTHOR REVIEWS HIS PREVIOUS RESEARCH INTO TECHNIQUES
TO ATTRACT CULTURALLY-DISADVANTAGED STUDENTS BACK TO THE
CLASSROOM AND TO MAINTAIN THEIR ATTENDANCE. A CONTINGENCY
SYSTEM WAS DEVELOPED TO IMPROVE PERFORMANCE, AND THE
VARIABLES SPEED AND ACCURACY WERE INVESTIGATED. THIS IS PAPER
NO. FOUR IN THE "VARIABLES INFLUENCING BEHAVIOR" PROJECT.
(LH)

ED015671

Report of the "Variables Influencing Behavior" Project

Dr. J. A. Jones, Director
Arizona State University
Tempe, Arizona

Paper No. 4

SUMMARY OF RESEARCH IN EXPERIMENTAL EDUCATION

Author:

Mark Laurence Berman
Department of Psychology
Arizona State University
Tempe, Arizona
August, 1967

U.S. DEPARTMENT OF HEALTH, EDUCATION & WELFARE
OFFICE OF EDUCATION

THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE
PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS
STATED DO NOT NECESSARILY REPRESENT OFFICIAL OFFICE OF EDUCATION
POSITION OR POLICY.

EM 005 629

Summary of Research in Experimental Education

Mark Laurence Berman
Department of Psychology
Arizona State University
Tempe, Arizona

INTRODUCTION

Over a span of three and a half years, an initial study in applied anthropology has progressively come into sharper focus as a research in the experimental manipulation of students' performance on educational materials. The control of human learning behavior demonstrated in the reported experimental findings (Berman. References # 1 to # 8) has direct application to the practical problem in acculturation with which we started.

The original study population was a group of young adult Yaqui Indian and Mexican-American fourth grade drop-outs. The problem was how to use operant techniques to attract these culturally disadvantaged men back to the classroom and maintain their attendance in educational programs.

With attendance and interest established, the problem changed to the management of contingencies in order to shape better performance on the programmed materials, and thereby change these students' past history of academic failure to a current experience of success. Stating the problem in terms of contingency management focused attention on two measures of successful performance on such materials, -- speed and accuracy. The program presented eighth grade level English grammar and syntax; frames were written with the students' cultural backgrounds in mind.

When the contingency system produced the desired results, the question arose

as to how fast and accurately, would students work if they were placed under appropriate reinforcement conditions. How stringent could demands be made before straining occurred? The research design developed to this point was adapted for use with a subject population of undergraduate college students.

Experiments were then devised to differentially reinforce speed and accuracy, and to provide data on preference, switching behavior, and total output. As a means of systematically increasing or decreasing demands made on the students, an adjusting schedule of criteria, with stated step-sizes, was put into effect.

From the data produced, it is possible to demonstrate the effect of some variables that exert control over speed and accuracy, such as the application of immediate reinforcement, the relative values of the pay-off matrix, and the criteria levels including the step-sizes by which one schedule is changed to another.

From the observed behavior of students throughout these experimental manipulations, it is possible to suggest that such operant procedures can be applied to a variety of educational settings and materials. Contingency management can have a direct bearing on the practical problems of creating a learning situation under which the culturally disadvantaged can be persuaded to accept further education, as well as one under which a college student can be persuaded to consistently perform at or near his maximum potential.

These first approximations to usable procedures suggest further research and a number of possible approaches. To give only one example, it is immediately apparent that a problem exists concerning when we should alter reinforcement contingencies and criteria levels for a particular subject, -- a complex decision that should take into account past and present performance of this student and of all earlier students, as well as the varying length and difficulty of individual sections of the programmed material. Perhaps only a computer could give an answer

based on adequate consideration of all the data.

This paper reviews and describes this extended research in experimental education. Full details of procedures and program structure can be found in the series of individual papers.

Shaping Attendance

The conditioning history of the subjects showed a paucity of positive reinforcers with a concomitantly large background of aversive control. Prior to exposing them to the proposed English program, the first step was to acquaint them with the University environment. With the co-operation of faculty and staff, students toured such areas as chemistry laboratories and zoological displays, listened to lectures, and visited the Student Union. The preliminary effect of these operations was more relaxed behavior on the part of the students.

After several weeks of this kind of experience, the students began inquiring into more intimate details of college life, particularly student activities related to study habits, lectures, and social life. In the following months their questions began to relate more to their own personal situation; they asked if it was possible for them to get into college, and if their children could do so.

After they had become familiar with their new environment, they began classroom work in arithmetic. Following completion of this program, a number of the students began the intermediate level English program. Class size was kept below seven students since it had been found that students performed better under this condition.

Maintaining Attendance

Working with programmed materials appeared to be sufficiently motivating in itself, and reinforcement for these culturally disadvantaged adult students

consisted solely of knowledge of correct results plus praise from their instructor. Reinforcers of extrinsic (non-monetary) value only were sufficient to maintain their participation in the program. It is to be noted, however, that in later studies with college students as subjects, other kinds of reinforcers (money) were necessary. It may be that college students are accustomed to a fairly high level of success, and may not work just to receive knowledge that their answers are correct.

Programmed Instruction and Culturally-Disadvantaged Students

Programmed instruction was found to be particularly effective with the Yaqui and Mexican-American students. Because of such factors as step-by-step organization of material, self-pacing, and immediate feed-back knowledge of correctness of answers, programmed instruction results in relatively high accuracy levels. During their grammar school years, these students reported that they had usually been incorrect on their test answers a high percentage of the time. By normal standards, their early performance on the programmed material in intermediate level English was quite poor. However, to them, even a 50% accuracy level was highly rewarding. Getting every other answer correct was a large improvement over their past performance in the schoolroom.

Their first reaction was that of surprise upon making so few errors, relative to past performance. Later, as their experience with the program increased, their error rate dropped drastically. The initial exclamations of surprise turned to pride at mastery of the programmed material. Accuracy levels reached as high as 98%, and the mean level reached was 94%.

In order to perform effectively on programmed materials the student must pick out key words and phrases in the frames. Careful reading is required. The intermediate level English program was designed so the student would begin working

with very simple English sentences and slowly be exposed to more and more complex sentence constructions. Likewise, vocabulary levels began very low and proceeded to rise systematically in difficulty. Students were provided with dictionaries and thesauri in order to (1) supply the definition of words and inform students of words with similar and opposed meanings, and (4) to introduce students to the use of these important reference sources.

The training implicit in this carried over into other areas. Observation of students' verbal responses and their performance on the program showed that not only did their spelling, punctuation, and related behaviors improve, but their verbal abilities showed a similar gain. Part of this improvement may have been due to mere practice effect, i.e., the opportunity or necessity for English conversation. Also, as the aversive nature of the educational setting wore off the students' progressive relaxation may have been a further cause. However, even with allowance for the two preceding factors, it appeared that working through the English program had positive effects on the verbal performance of the students. For example, vocabulary and sentence structure became progressively more sophisticated and the rate of improper grammatical usages dropped, in some cases dramatically.

Reading Comprehension Program

Following completion of the intermediate level English program four of the seven students involved in the program were given another program specifically designed to shape reading comprehension. The results and the procedures involved are fully discussed in Berman (August, 1966). Content of the frames was related to homonym, synonym, and antonym distinctions, vocabulary building, discrimination of key phrases in successively longer paragraphs, etc. Accuracy levels ranged from 92.0 to 98.6%.

Objective Testing Measures

A battery of objective tests were administered to students following completion of both the intermediate level English program and the reading comprehension program. Each of the tests had two forms. The first form listed was administered after the intermediate English program, the second form after the reading comprehension program (to the same students). The tests used were:

- (1) Otis Group Intelligence Test, Advanced Examination, Forms B and A
- (2) Ohio Scholarship Tests, Techniques in Reading Comprehension for Junior High School, Grade 9, First and Second Every Pupil Tests
- (3) Ohio Scholarship Tests, Spelling and Vocabulary, Grade 6, First and Second Every Pupil Tests
- (4) Ohio Scholarship Tests, Elementary Reading-General Ability, Grade 6, First and Second Every Pupil Tests
- (5) Durrell Sullivan Reading Achievement Test, Word Meaning and Paragraph Meaning, Intermediate Forms B and A
- (6) Stanford Achievement Test, Spelling and Language, Forms W and X
- (7) Metropolitan Achievement Tests, Intermediate Reading Test, Word Knowledge and Reading (Comprehension), Forms Bm and Am.

Results of Tests with Time Limits

Berman (August, 1966) describes the results of the testing. In brief review, the required time limits seriously biased these tests against students who could not read test questions fast enough, or who worked through material at a deliberate pace. The bias was found to be of different weight and consequence for different tests. It was pointed out that a critical factor in diagnosing a student's facility in the use of English was the nature of the particular test administered. For example, the mean score on the Metropolitan Achievement Tests (Reading) was 8th grade plus, whereas the mean score on a related test, the Stanford Achievement Test in Spelling and Language was below the 5th grade level.

Results of Tests Without Time Limits

The first administration of the tests was accompanied by adherence to the time limits called for by the test publishers. On administration of the second form of the tests, students were allowed to work to completion after following the required protocol. Very large performance increments were noted on every test. For example, the mean score on the Ohio Scholarship Test, Reading Comprehension, on the second administration with time limit in effect was in the 2nd percentile for 9th grade students. Allowed to work to completion, the mean percentile increased to 39.3%. For the Ohio Elementary Reading-General Ability Test, scores increased from a mean of the 27th percentile for Grade 6 to a mean of the 48.5 percentile. As a final example, Reading scores on the Metropolitan Achievement Tests increased from a grade equivalent of 5.36 with a time limit imposed on the second administration, to a grade equivalent of 9.0 when allowed to work to completion.

On the Otis Intelligence Test, it was found that most questions were missed with time limits imposed because of lack of time, rather than because they were too "difficult" for the students to comprehend. The results of allowing students to work to completion are as follows: Mean IQ within time limits was 85.33, or in the bottom 11.3% of the test population. Mean IQ to completion was 106, or in the top 31.17% of the test population. Allowing students to work to completion raised the mean IQ over 20 points.

In general, students did not show much improvement in test scores when comparing the results of the first administration (time limits imposed) to the second administration (time limits also in effect). Probably the particular tests administered were not sensitive to information gained on the reading comprehension program. It would have been better to have administered one form of each of the

tests prior to the start of the intermediate level English program and the other form of the tests after completion of this program.

THE CONTINGENCY SYSTEM

Because of the huge background deficits possessed by the Yaqui and Mexican-American students it was clearly insufficient to be concerned merely with inculcating certain behaviors (such as facility in the English language). If these students were to acquire mastery of enough subject areas and to attain high school equivalency certificates, it was necessary to train them much faster than previously. The procedures required for carrying out the rapid shaping of knowledge acquisition would have to take into account the students' very low-level baseline behaviors. An adjusting schedule technique, first described in Berman (#5, August, 1967) was developed.

Advantages of an Adjusting Schedule

The adjusting technique is particularly effective with the so-called culturally-disadvantaged student, since he usually begins working on a program with a poorly-developed repertoire. The use of adjusting schedules allows for shaping poor entering behavior towards better and better performance.

An adjusting schedule permits reinforcement criteria to change continually as student performance levels change over time. By making reinforcement contingent upon progressively higher performance levels, it is possible to shape behavior in the direction of greater speed and accuracy. If fixed reinforcement schedules are used, there can be no compensation for change in the subjects' performance levels nor allowance for individual differences in performance. Also, since material varies in difficulty as the program progresses, the use of adjusting schedules helps to avoid setting performance criteria too high or too low in relation to the

difficulty of the particular section of the program currently being worked on.

Required Changes in Research Design

In the first application of the adjusting schedule technique (with the Yaqui and Mexican-American students) no baseline measures were gathered to determine the students' "normal" accuracy and speed on the programmed materials. The reinforcement criteria or contingencies were arbitrarily set for all students at the same levels. Consequently, for the first quarter of the program (a commercially-prepared program, Programmed English, by M. W. Sullivan. New York: MacMillan, 1963) the students had no trouble meeting both time and error criteria. For the remainder of the program the contingencies "took hold," i.e., the students were subject to more stringent constraints in regard to the range in which performance levels could fluctuate and still lead to reinforcement (in the nature of points of no intrinsic value).

For all practical purposes performance on the first quarter of the program was under poor control. Because the initial criteria were not based on baseline measures, meeting them was much too easy. The problem of where to set initial contingencies is not solved simply by gathering baseline information on the students involved. It has been found (see Berman, July, 1967 - Simultaneous Maximization...) that setting initial reinforcement criteria at the average time and/or accuracy scores made on the baseline sections of a program usually results in initial criteria which are too lenient. A better procedure would be, first, to set initial criteria at, say, 67% of the baseline values, and, second, to utilize the performance of previous subjects on the same program to aid in establishing initial criteria (discussed in Berman, July, 1967 - Simultaneous Maximization...).

Adjusting Schedule with Counter-Balanced Design

A second study using the adjusting schedule technique was carried out with college students as subjects. Several students formed a control group which completed the program (Daniel P. Kimble. Physiological Psychology, Reading, Massachusetts: Addison-Wesley Publishing Company, 1963) with no criteria in effect at any time. The experimental subjects were exposed to a counterbalanced procedure, in which periods of no criteria were alternated with periods when time criteria were in effect.

Berman (October, 1967), discussed the effects of the use of the counter-balanced design to provide more precise evaluation of the effects of the adjusting schedule technique. The effects of changing from no criteria to a time criteria (and vice-versa) were clear. When the time criterion was in effect speed increased significantly, with a decrease in accuracy, over the previous no-criteria portion of the program. When the no-criteria condition was established (or re-established) speed decreased significantly and accuracy increased. As in the first adjusting schedule study, points with extrinsic value only were used as reinforcers. It was found that, in contrast to the first study, in which the programs used were directly related to the interests of the students, college students would terminate their participation unless the programs were directly related to their major field or interests.

Simultaneous Maximization of Speed and Accuracy

Previously, the effects of time contingencies on speed and of error contingencies on accuracy of performance had been studied separately: in some cases speed was shaped, in other cases accuracy was shaped. Of practical importance was a method for shaping both speed and accuracy simultaneously. The details of this study are described in Berman (July, 1967 - Simultaneous Maximization...).

The results of the study demonstrate that for some programs, time taken to complete can be reduced to at least an average of 38% of the time the publisher of the program says was required by students in his pre-publication test population. (Note: the program used in this study was written by Cynthia D. Buchanan -- A Programmed Introduction to Linguistics: Phonetics and Phonemics, Boston: D. C. Heath and Company, 1963). The mean accuracy level of our three subjects was 97.5%.

It is believed that through the use of more refined procedures speed on the program could be reduced to an average of approximately 25% of the time required by the pre-publication test population, while maintaining accuracy at a high level.

Student Preferences for Speed Versus Accuracy Contingencies

A study was carried out to investigate the relative preferences of students for working under speed or accuracy criteria. A choice situation was in effect, in which students chose to meet either the time or the error criterion in effect in each section of the Buchanan program. Initial criteria were individually determined for each student, as the result of performance on no-criteria baseline sections of the program.

Definite preferences (i.e., 87% of the choices) were established for working under accuracy conditions, even though students had greater success at meeting the reinforcement contingencies when they chose to work rapidly rather than accurately. Several hypotheses were advanced to account for the choices. They are discussed in detail in Berman (July, 1967 -- The Effect of Response Cost on Error Versus Time Choices...).

Compared to the situation in which accuracy and speed criteria have to be

met simultaneously in order to receive reinforcement, both accuracy and speed are poorer when a choice situation exists. The explanation offered is that in a choice situation the student is able to neglect one dimension of performance while concentrating on another dimension. Though it is far more difficult to attempt to be accurate following a number of consecutive choices of the time contingency than to attempt to work quickly following a number of consecutive choices of the accuracy contingency, in each case too much emphasis on one contingency hurts the student's chances of meeting the other contingency. If a student consistently chooses the speed criterion through the first half of a program, it would probably be next to impossible for him to switch his emphasis to attempting to meet the accuracy criterion in the second half of the program and be successful more than a fraction of the time. This is a function of the cumulative nature of most programs, in which mastery of earlier information is a prerequisite for high-level accuracy on later material. With both speed and accuracy required at the same time the student cannot let one performance dimension deteriorate to the point of no return, without forfeiting all possibility of receiving reinforcement later in the program.

Response-Cost Procedure

A second study of student preferences for working under speed or accuracy criteria was carried out using a response-cost procedure. This study is described in Berman (July, 1967 -- The Effect of Response Cost...). With response cost in effect subjects who attempted to meet the error criterion and were unsuccessful, had 1, 5, or 10 cents deducted from money previously earned in this program. Under this condition, subjects showed marked preferences for time choices. Whereas error choices had occurred on an average of 87% in the previous no-cost study, with response cost in effect the average error choice was only 37%.

While the response cost system led to significantly fewer error choices, a negative side effect was seen in a lower accuracy level than that which prevailed in the no-cost choice study. In the latter mean accuracy was 95%, while the response-cost subjects performed at a mean accuracy level of 85%, with a range from 79% to 89%. In addition, response cost led to completion times one hour longer on the average than for the no-cost study (9 and one-quarter as compared to 8 and one-quarter hours). It was also found that the greater the response cost the greater the variability in the number of errors made on a section-by-section basis. No other systematic differences between 1, 5, and 10 cent response-cost groups were found.

The Effect of Adjustment of Step-Size on Performance

The initial preference study was also concerned with the effects of the magnitude of the adjustments or step-sizes used in the adjusting schedules. The step-sizes used were 1, 2, and 3 minutes and 1, 2, and 3 errors. During the first half of the Buchanan linguistics program each group had a particular set of step-sizes in effect for time and error criteria. For example, students in Group 1 had 1 error and 1 minute step-sizes. During the second half of the program the step-sizes for Group 1 changed to 3 errors and 1 minute.

The smaller the error step-size on the first half of the program the greater the preference was to choose accuracy rather than speed. There was also a positive relationship between the step-size and the total number of switches in the program. The smaller the step-size the less switching. When the criteria adjust in small steps it is easier to maintain control over speed and accuracy. This relates to the finding of the response-cost study that the larger the response cost the more variable are the number of errors made per section.

The effect of the changed step-sizes on the no response-cost preference study subjects was almost undetectable. Subjects continued to attempt to meet the error criteria. On the response-cost study the effects of changing step-sizes were marked. The larger cost subjects (i.e., those with 5 cent response cost) showed the greatest changes in preferences. No subjects on 10-cent cost completed the program, perhaps due to the very aversive nature of the large cost. The 1-cent cost subject showed very little change in preference following step-size change. This is in keeping with the almost total lack of change found in the no-cost subjects.

When error step-sizes changed on the second half of the program for the response-cost subjects (from 1 minute-1 error step-sizes to 1 minute-3 error step-sizes), successful error choices increased from 48% to 66%. This is another example, in keeping with the results of the earlier preference study, of the fact that improvements in success at meeting error criteria do not produce concomitant increases in the number of times error criteria are chosen. Berman (July, 1967 -- Simultaneous Maximization...) discusses the possible reasons responsible for this phenomenon.

The effects of larger step-sizes (during the second half of the program) for errors was not noticeable during the no-cost study since the subjects were already choosing error criteria consistently, and no further increase was possible. The response-cost procedure is a more sensitive device for measuring the effects of step-size changes on relative preferences for time versus accuracy.

Epilogue

The initial research described in this paper began with the development of techniques to (1) attract culturally-disadvantaged students back to the classroom setting and (2) to maintain their attendance in the educational programs. Later

a contingency system was developed to shape better performance, as opposed to simply maintaining their interest. The contingency system was first tried out on Yaqui and Mexican-American students and then further developed and adapted for use with college students.

While this latter work was underway the preference studies were carried out to determine how fast and/or accurately students would work under various reinforcement contingencies. It was desired to get information as to how stringent contingency requirements could be made before straining occurred. The experiments on time versus error choices under no-cost conditions supplied data of this kind.

The response-cost study was carried out to determine the effects of cost on preference behavior. The step-size manipulations which occurred during the first preference study yielded information concerning the effects of step-sizes on performance. Finally, a study was carried out for the purpose of developing a technique to shape greater speed and accuracy simultaneously, a potentially valuable tool for classroom use.

In conclusion, the work described in this paper demonstrated that many aspects of student performance are subject to improvement through the application of proper control and shaping procedures. In addition, an analysis of performance (as distinct from the shaping of performance) was carried out, in order to isolate some factors potentially critical in human learning situations.

* * * * *

ACKNOWLEDGEMENT

The author would like to take this opportunity to publicly thank Mrs. Helen Sherer for her invaluable editorial assistance on this and several earlier reports (i.e., V.I.B. Reports # 2 and 3).

REFERENCES

1. Berman, M. L. The development of proficiency in the use of the English language in groups of non-Western indigenous peoples through programmed instruction. May, 1964. Tempe, Arizona: Department of Anthropology, Arizona State University. (Variables Influencing Behavior Project, Dr. J. A. Jones, Director, Paper No. 1) 151 pp.
2. Berman, M. L. Some considerations in the education of indigenous groups in the southwest. Santa Monica, California: System Development Corporation. 19 July, 1965. Professional Paper SP-2148. 11 pp.
3. Berman, M. L. Experimental explorations in programmed instruction and objective testing measures. Tempe, Arizona: Department of Anthropology, Arizona State University (V.I.B. Project, Paper No. 2) August, 1966. 43 pp.
4. Berman, M. L. Towards greater extension of psychological principles to programmed instruction: the use of adjusting schedules of reinforcement. National Society for Programmed Instruction Journal (NSPI), 6, no. 1, 1967, January, 7-9.
5. Berman, M. L. Shaping performance on programmed materials. Psychological Reports, 1967, 21, August, 29-32.
6. Berman, M. L. Experimental analysis of performance on programmed materials. National Society for Programmed Instruction Journal (NSPI), 6, no. 8, 1967, October.
7. Berman, M. L. Simultaneous maximization of speed and accuracy on programmed materials through the use of adjusting conjunctive schedules of reinforcement. Submitted to NSPI Journal, July 12, 1967.
8. Berman, M. L. The effects of response cost on time versus error choices on programmed materials. Tempe, Arizona: Department of Anthropology, Arizona State University. (V.I.B. Project, Paper No. 3) July, 1967.

REFERENCES

Continued

9. Buchanan, Cynthia D. A Programmed Introduction to Linguistics: Phonetics and Phonemics. Boston: D. C. Heath and Company, 1963.
10. Kimble, Daniel P. Physiological Psychology. Reading, Mass.: Addison-Wesley Publishing Company, 1963.
11. Sullivan, M. W. Programmed English. New York: MacMillan, 1963.